DOCUMENT RESUME

ED 288 392 FL 017 032

AUTHOR Griffiths, Patrick
TITLE Constituent Structure in Text-Copying.
PUB DATE Nov 86
NOTE 44p.; In: York Papers in Linguistics No. 12, 1986; see FL 017 027.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Processes; Comparative Analysis; Comprehension; Foreign Countries; Language Processing; Language Tests; *Linguistic Theory; *Phrase Structure; Psycholinguistics; Research Methodology; Semantics; *Sentence Structure; *Structural Analysis (Linguistics); Suprasegmentals; *Syntax; *Transformational Generative Grammar

ABSTRACT

A study explored the use of a text-copying task for the determination of syntactic constituent structure. It was predicted that the task would be a naturalistic, reasonably direct, and sensitive psycholinguistic research method. In two experiments, 70 subjects wrote out copies of typed passages. The points where they paused and looked back to the typed originals were recorded and aggregated across all subjects copying a given passage. The passages differed in content but contained sentences with comparable structure. In general, patterns of pausing were highly correlated across sentences that matched in structure, but reliably distinct in relevant control comparisons. Sentences, clauses, and many details of their internal constituency emerge clearly. These details are presented and compared with descriptions from the field of theoretical linguistics. Some suggestive evidence in support of two-stage parsing models is noted. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED288392

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S. Harlow

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CONSTITUENT STRUCTURE IN TEXT-COPYING[1]

Patrick Griffiths

(Department of Language
and Linguistic Science)

Abstract

In this paper a text-copying task is explored as a naturalistic, reasonably direct, and sensitive psycholinguistic method for the determination of syntactic constituent structure. In two experiments, 70 subjects wrote out copies of typed passages. The points where they paused and looked back to the typed originals were recorded and aggregated across all subjects copying a given passage. The passages differed in content but contained sentences with comparable structure. In general, patterns of pausing were highly correlated across sentences which matched in structure, but reliably distinct in relevant control comparisons. Sentences, clauses and many details of their internal constituency emerge clearly. These details are presented and compared with descriptions from the field of theoretical linguistics. Some suggestive evidence in support of two-stage parsing models is noted.

Introduction

It is a commonplace of linguistics that sentences are more than just concatenations of words. They have hierarchical constituent structure: words are grouped into phrases and these, in turn, may form other phrases or clauses, which finally make up sentences. Zwicky (1978) provides a valuable conspectus of the reasons which have led theoretical linguists to accept constituents intermediate between the word (or morpheme) and the sentence. Essentially, the recognition of constituent structure facilitates - and in some cases appears to be an indispensable basis for - generalizations about syntactic, semantic and even some phonological patterns. Psycholinguistic research, some of which enters Zwicky's survey and some of it to be mentioned below, has established that hierarchically-organized constituents also play a part in the mental processes that subserve human language use.

If it could be demonstrated that, in studies of sentence constituency, psycholinguists and theoretical linguists are looking at fundamentally the same phenomena (though from different angles) then the latters' accounts of syntactic structure could be useful to psycholinguists trying to understand such processes as sentence comprehension. If, furthermore, psycholinguists can develop

FL017032

BEST COPY AVAILABLE

York Papers in Linguistics 12 (1986) 75-116
© the author.

good ways of revealing the sentence constituents that language users actually employ in listening, reading, writing and talking, then their results could be a worthwhile addition to the empirical data-base of theoretical linguistics.

An ideal psycholinguistic method for probing the constituency of sentences should be strong on the following three characteristics: task naturalness, simultaneity of subject's language processing and experimenter's observation, and sensitivity to details of constituent structure. The first two requirements relate to validity: findings from highly artificial tasks cannot confidently be generalized to normal language use, and time elapsing between processing and observation allows scope for other variables to influence the results and thereby complicates their interpretation (the second desideratum is but a slight extension of Levelt's (1978:4) notion of "simultaneous measurement"). Sensitivity to detail is needed partly to enhance reliability, but also to enable one to gather facts beyond those which are anyhow obvious.

Levelt (1969, 1970) has found ways of eliciting judgements of the relative strength of the relationships between all possible pairs of words in a sentence. Hierarchical cluster analysis techniques translate matrices of these judgements into tree diagrams. The method appears to be sensitive, but the judgemental task is not one that figures in everyday language use and the subjects' judgements do not coincide in time with their initial processing of the sentences. Excepting perhaps for naturalness, the same points can be made about Johnson's (1965) analysis of transitional error probabilities: subjects memorized sentences and their recall protocols were examined to determine, across a group of subjects, the probability of error in each transition between the successive words of a given sentence. Martin (1970:160) simply asked subjects to "... arrange the words of each sentence into natural groups ... into whichever and however many parts seem intuitively natural". The nature of the resulting cluster analysis trees speaks for the sensitivity of the method, but again the judgements followed whatever processing the subjects did on first exposure to the sentences and the judgements were, as in Levelt's (1969, 1970) work, metalinguistic ones that language users need never have ordinary occasion to make overtly. The approach is also so transparent as to amount to nothing more than a systematic collation of opinions.

The adaptation of subjective click location tasks by Bever and his associates for the psycholinguistic investigation of syntactic constituency is well known (see Fodor, Bever & Garrett, 1974, for a synopsis). Subjects hear a sentence with a superimposed click and afterwards (i.e. the method is not a simultaneous observation one)

indicate where in the sentence they thought they heard the click. The clicks are held to drift towards - and/or to be more accurately located when they occur in - certain types of inter-constituent boundary. The method has been cogently criticized by Larsen (1971), and Olson & Clark (1976) also raise important questions about the interpretation of click experiment results. Here it will suffice to add that the task is not a natural one and it appears, from Bever, Lackner & Kirk's (1969) study, to be somewhat insensitive to constituent structure below clause-level.[2]

The inferring of constituency from hesitation pauses in spontaneous speech (see Hawkins, 1971, for an example of this genre) and from pausing when subjects read aloud (as exemplified by the research of Grosjean and his collaborators, e.g. Grosjean, Grosjean & Lane, 1979) strike me as more promising directions. Speaking and reading are natural tasks (the latter, of course, only in literate communities) and simultaneity is assured by the fact that pausing is part of the performance. The work of Grosjean et al suggests that their method is sensitive to details of sentence structure. Hesitation pause analyses, such as Hawkins', seem less sensitive to clause-internal structure; perhaps because of the common practice of ignoring very short pauses (e.g. below 250 or 300msec, on the grounds that some of these may be merely stop consonant closures). For present purposes, the data from hesitations in spontaneous speech have two further drawbacks: some of the pausing may be occasioned by the speaker hunting for content words (rather than by the category of constituent being readied for output) and, secondly, the investigator cannot choose exactly which sentences the speakers will utter. Reading aloud eliminates these two problems, since the investigator specifies what the subjects are to read. There is a residual possibility, however, that both types of pause study may include some pause time which is not germane to the concerns of this paper: namely, it is conceivable that pause duration in speech could have other signalling functions besides the demarcation of constituents (see Butterworth, 1980:157).

The methods enumerated above are by no means the only ones that have been employed, but they are a representative selection. None of them yields constituent structure analyses which exactly match the clear cases over which there is a reasonable concensus within theoretical linguistics. Of course, the theoretical linguists might be mistaken, but it might alternatively be the case that the currently available psycholinguistic methods are too much subject to random errors and confounding factors. The method of text-copying presented in this paper involves observation that is nearly concurrent with the processing, shows considerable sensitivity to detail and, by and large, gives hierarchical constituent structures closer to those of theoretical linguistics than are produced by any other

technique known to me. It is also (see footnote 2) an economical method. For literate subjects the task is a reasonably naturalistic one. It amounts to doing something which is occasionally done anyway during the normal course of one's life as a language user. The processing constituents derived from data collected in this way are thus likely to have ecological validity and are unlikely to represent merely ad hoc strategies resorted to for coping in a novel experimental situation.

The first of the two experiments described below (immediately following an account of the rationale for the text-copying method) was mainly a methodological exploration. The second experiment was an attempt to use the method to examine particular structures which are of some interest to theoretical linguists.

Rationale[3]

Subjects were asked to write out copies of typed passages (just as one might copy out a quote from an academic article). The typed originals and the sheets of paper on which subjects made their copies were placed sufficiently far apart to prevent subjects from reading the original and watching their pens at the same time. In these circumstances, subjects who choose to guide their writing visually (and, fortunately, none of those in the two experiments reported below chose to write completely blind of visual control) are obliged to look at the typed original and take a stretch of it into short-term memory, then to write down that stretch from memory before looking back at the original for another stretch to load into short-term memory, etc. In principle, these stretches of material could be of, at least, the following two kinds:

> They could be linguistic constituents of whatever size the subject found convenient.
>
> Or they could be arbitrary strings of words, e.g. one word at a time or five words at a time.

The motivating hypothesis for the two studies presented here is that subjects will generally follow the first strategy.

If different subjects copying the same sentence transfer constituents of different sizes, their aggregated results will reveal the hierarchical structure of the sentence. To see why this should be so, consider a two-clause sentence in which each clause has two short phrases as its intermediate constituents:

$$S[_{Cl}[\text{PHRASE PHRASE}]_{Cl}[\text{PHRASE PHRASE}]].$$

Imagine that it is copied by three subjects who differ in short-term memory capacity (or preference or perhaps even

just in rashness): Subject A copies a whole sentence at a time, Subject B copies a clause at a time and Subject C copies only a phrase at a time. For these three subjects the aggregated record of look-back-for-another-chunk-to-copy pauses will be (with the subscript letters showing where each subject paused):

$$\text{PHRASE}_{C}\text{PHRASE}_{BC}\text{PHRASE}_{C}\text{PHRASE}_{ABC}$$

It is clear that the number of pauses in each locus will rather directly reflect the constituent structure of the sentence; and there is no reason why, with more and varied copiers, the internal structure of phrases should not be revealed too.

There is a straightforward algorithm for converting numbers of pauses made by a group of people copying a sentence into a bracketing which represents the hierarchical constituent structure which that group imposed on that sentence. It is illustrated by means of some data collected in Experiment 2. Forty-seven subjects copying a set of five-word sentences yielded the following pattern of pauses:

$$\text{ARTICLE}_{2}\text{NOUN}_{9}\text{MODAL}_{3}\text{BE}_{8}\text{PAST-PARTICIPLE}_{40}$$

(that is, 2 of the subjects paused after the article, 9 after the noun, ... and 40 of the 47 stopped writing and looked back to the original text when they came to the end of the sentence). The procedure[4] is: (a) Find the smallest cluster of pauses (in this case it is that of the two people who paused after the article) and bracket together into a single constituent the constituents which immediately flank that cluster. (b) Delete that cluster of pauses and return to step a. For the specimen data, successive passes through steps a and b give:

$$[\text{ARTICLE NOUN}]_{9}\ \text{MODAL}_{3}\ \text{BE}_{8}\ \text{PAST-PARTICIPLE}_{40}$$

$$[\text{ARTICLE NOUN}]_{9}\ [\text{MODAL BE}]_{8}\ \text{PAST-PARTICIPLE}_{40}$$

$$[\text{ARTICLE NOUN}]_{9}[[\text{MODAL BE}]\ \text{PAST-PARTICIPLE}]_{40}$$

$$[[\text{ARTICLE NOUN}]\ [[\text{MODAL BE}]\ \text{PAST-PARTICIPLE}]]_{40}$$

This bracketing is of course equivalent to the tree diagram in Figure 1, where the clusters of pauses have been resuscitated to show the relationship between them and the resulting tree.[5] To a linguist this tree is an attractive invitation to deploy the labels S, NP, VP and AUX.[6] The algorithm, however, does not provide the labels; that is a matter for subsequent interpretation.

As it has been stated thus far, the algorithm makes no allowance for ties. A rider added to the end of step a will rectify this: "If two or more clusters of pauses tie for the

position of smallest cluster, deal with them simultaneously". Thus $T_5U_2V_7W_2X_2Y_2Z_{10}$ will be mapped by a single pass through the procedure into:

$T_5\ [U\ V]_7\ [W\ X\ Y\ Z]_{10}$.

The only point that remains to be made about the rationale for the text-copying method is that effects attributable to sentence structure can be disconfounded from the pause-inducing propensities, if any, of particular words by amalgamating results from two or more sentences which are structurally comparable but contain different words.

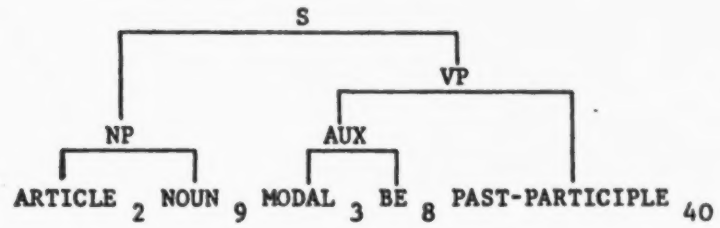


Figure1. Tree diagram based on the amalgamated pause distribution of a set of four 5-word sentences used in Experiment 2.

Experiment 1

Subjects

Forty-six senior undergraduate and beginning graduate Linguistics students participated in the experiment in fulfilment of a course requirement. Psycholinguistic investigations of constituent structure had not yet been discussed in the course at the point when the experiment was conducted. Half of the group, 23, acted as subjects and the remaining 23 were observers.

Materials

Four typed 62-word passages (see Figure 2, opposite) were used. The COHERENT 'SEMANTICS' and COHERENT 'PHONETICS' passages were loosely derived from introductory textbooks (and were therefore expected to strike the Linguistics student subjects as reasonably familiar and not particularly abstruse). The two JUMBLED passages are variants of the corresponding COHERENT passages with the words randomly re-ordered. Lineation and the minimal punctuation were exactly as shown in Figure 2.

Each of the COHERENT passages contains four sentences. They were designed to correspond in their syntactic structure across passages. The first sentence in the COHERENT 'SEMANTICS' passage (henceforth, sentence S_1) is 14 words long and has a clause 'extraposed' from subject position. This was matched by the fourth sentence in the COHERENT 'PHONETICS' passage, P_4, which is also 14 words in length and also has a clause 'extraposed' from subject position. The second COHERENT 'SEMANTICS' sentence, S_2, is

COHERENT 'SEMANTICS'

It is common knowledge that children will often misunderstand the meaning of a word. In the vast majority of cases such misconceptions will be corrected by later experience. If this does not happen a semantic change will take place in the usage of the new generation. A further factor which may lead to changes in meaning is the loss of formal motivation.

COHERENT 'PHONETICS'

A notable problem which has made vowels difficult to transcribe is the absence of articulatory contact. When the consonants are formed a trangible stricture is usually present between the surfaces of the two articulators. More than half of the time this contact can be sensed by the investigator. It is generally recognized that phoneticians may sometimes misclassify the quality of a vowel.

JUMBLED 'SEMANTICS'

If experience this later does by not corrected happen be lead that to a will semantic misconceptions change such will cases knowledge changes take of place majority in vast the the usage in of common in is meaning word the a new of generation meaning a the further it is motivation of misunderstand factor often which will may children formal loss the.

JUMBLED 'PHONETICS'

The between surfaces present of usually the is two stricture phoneticians articulators tangible more a than formed half are of consonants problem the true time when this contact articulatory can misclassify a vowel of be absence sensed the by is the transcribe investigator to quality it difficult is vowels generally made recognized has that which the phoneticians problem may notable sometimes a.

Figure 2. The four passages used in Experiment 1. Excepting for the headings, which did not appear on the versions actually used, the format of each passage is shown exactly as it was presented to the subjects. Two errors that appeared in the originals are retained, uncorrected, here: tangible was incorrectly typed as trangible in the COHERENT 'PHONETICS' passage and the word phoneticians appears once too often in the JUMBLED 'PHONETICS' passage.

likewise equal in length (14 words) and comparable in structure (passive with a preposed adverbial phrase) to the penultimate COHERENT 'PHONETICS' sentence, P_3. S_3 and P_2 are both 18 words long and have two-clause subordinating conjunction structures. S_4 and P_1 match in having 16 words each and in having relative clauses on their subject NPs.

For purposes of analysis, the JUMBLED passages are treated as if they were each composed of four sentence counterparts equal in length to the corresponding parts of the relevant COHERENT passage. Thus (using prime marks to denote these random strings of words) the 16-word sentence at the start of the COHERENT 'PHONETICS' passage, P_1, has as its control counterpart P'_1, the first 16 words of the JUMBLED 'PHONETICS' passage ("The between ... than"); and so forth.

Procedure

Testing took place in two separate sessions (with any given subject attending only one of the sessions) in a language laboratory. Each subject made written copies of two of the typed passages. An arbitrarily selected half of the subjects at each testing session copied the COHERENT 'SEMANTICS' and JUMBLED 'PHONETICS' passages. The remaining subjects copied the COHERENT 'PHONETICS' and JUMBLED 'SEMANTICS' passages. Half of each of these sub-groups copied their COHERENT passage first and half did their JUMBLED passage first.

The experimental lay-out for right-handed subjects was as follows. In front of the subject was a blank sheet of paper on which the copy was to be written and a few inches to the right of the subject's right elbow was the typed passage to be copied. This arrangement necessitated a distinct swivel of the subjects' heads when they looked up from their writing to scan the original from which they were copying. To the left of the subject sat the observer. Observers had before them identical typed copies of the particular passages which their subjects were copying. On these, observers simply made a mark after the last thing their subject had written each time the subject stopped writing and glanced back at the original passage on the right. For left-handed subjects the lay-out was reversed: the observer sat on the subject's right and the original passage was placed to the left of the subject's left elbow.

Subjects and observers heard separate tape recorded instructions through language laboratory headsets immediately before the start of testing. Subjects were asked to make "quick and accurate" written copies of the typed passages. The instruction was reiterated as "aim for accuracy but work reasonably quickly too". They were told that their individual observers would be watching them writing and would be "noting something about the way you

write". They were asked not to move the typed passages away from the position beyond their writing elbows.

The task of noting the points in the passages at which subjects stopped writing and looked back at the original across their writing arms was explained to observers. They were told to keep the original passages covered until they and their subjects were ready to start. They were instructed to allow a rest of approximately one minute between the copying of the first and the second passage, and to substitute the second passage - covered until copying was about to start - for the first during this interval. For observers the purpose of the experiment was characterised as an attempt to "find out where people stop when they are making a written copy of a text". They were offered no further information regarding the hypothesis under investigation. Both subjects and observers were promised that a fuller explanation would be given at the end of the experiment (and this was duly delivered).

Results

Half of the subjects had copied a COHERENT passage before copying a JUMBLED one and half of them had copied JUMBLED before COHERENT. Inspection of their separate results revealed only one striking difference: the 5 subjects who copied JUMBLED 'SEMANTICS' as their first passage paused on average 27.20 times in dealing with this particular passage, whereas the 7 who had it as their second passage yielded a mean for this passage of 19.43 pauses. There is no ready explanation for this difference and for the other three passages order differences appear to be negligible. Order effects were therefore ignored and data

n	COHERENT 'SEMANTICS'		JUMBLED 'PHONETICS'	
11	mean	11.18	mean	18.73
	sd	5.71	sd	8.83

n	COHERENT 'PHONETICS'		JUMBLED 'SEMANTICS'	
12	mean	17.17	mean	22.67
	sd	6.70	sd	6.91

Table 1. Means and standard deviations of pauses per subject for the four passages of Experiment 1. Passage-final pauses, where everyone had of necessity to stop writing, are not counted here and are omitted from all subsequent analyses.

from subjects copying a given passage either first or second were combined to produce the means and standard deviations given in Table 1.

Subjects looked back to consult the original text more often when copying JUMBLED passages than when copying COHERENT passages (one-tailed paired-observation t-tests give $t=3.647$ ($df=10$, $p<.005$) for COHERENT 'SEMANTICS' versus JUMBLED 'PHONETICS' and $t=2.312$ ($df=11$, $p<.025$) for COHERENT 'PHONETICS' versus JUMBLED 'SEMANTICS').

These differences arise from subjects carrying across sequences of just over four words long (mean = 4.05) from eye to pen before each pause for COHERENT texts, but fewer than three words (mean = 2.85) each time for the JUMBLED texts. Only 5 of the 23 subjects defied this trend and paused more often in their COHERENT passages; in each of these cases the difference was small and 4 of these five reversals occurred when JUMBLED 'SEMANTICS' followed COHERENT 'PHONETICS'.

Although the two JUMBLED passages do not differ significantly in the average numbers of pauses which they attracted (two-tailed independent-sample $t=1.198$, $df=21$, $.30>p>.20$), the COHERENT 'PHONETICS' passage was clearly more difficult than the COHERENT 'SEMANTICS' passage (two-tailed independent-sample $t=2.296$, $df=21$, $p<.05$). A factor which might contribute to this is that the 'PHONETICS' passages contained more long words. Note in Figure 2 that the 'PHONETICS' passages extend over six lines, versus five lines for the 'SEMANTICS' passages, despite their having the same number of words as the 'SEMANTICS' passages. This is highlighted by pauses that fell <u>within</u> words: there were only 4 intra-word pauses, in total, for the 'SEMANTICS' passages (all of them in the JUMBLED version), but the 'PHONETICS' passages accumulated a total of 23 such 'spelling check' pauses (8 for the COHERENT version and 15 for the JUMBLED version). In compiling Table 1 and for the purposes of all subsequent analyses, a pause within a word was treated as if it had occured <u>after</u> that word (and ignored altogether if that subject had also paused after the word). Compared to the total of more than 800 pauses, the number of intra-word pauses is in any case trivially small.

The JUMBLED versions of the passages were intended to provide between-subjects control comparisons for any syntactic effects on pausing in the corresponding COHERENT passages. The mean of 11.18 pauses for the COHERENT 'SEMANTICS' passage is indeed highly significantly different from the figure of 22.67 for JUMBLED 'SEMANTICS' (one-tailed independent-sample $t=4.324$, $df=21$, $p<.005$), but the difference between the two 'PHONETICS' passages fails to reach significance (one-tailed independent-sample $t=0.480$, $df=21$, $.35>p>.30$). The differences in difficulty between the 'PHONETICS' and the 'SEMANTICS' passages catalogued in

this and the preceding paragraph would have provided an excuse had syntactic effects attributable to parallels between the sentences of the two COHERENT passages failed to manifest themselves. However, the correlations discussed below largely obviate the need for excuses.

The minimal syntactic effect that one might expect would be large concentrations of pauses at the ends of sentences in the COHERENT passages. (The end of the fourth sentence in each passage, of course, corresponds to the end of the whole passage. Since all subjects had to stop when they came to the end of a passage, this expectation can only be tested for the first three sentences in the COHERENT passages.) Table 2 shows that the subjects overwhelmingly paused to look back at the original text when they reached the full stops at the ends of the first three sentences. It could perhaps be argued that a full stop followed by a double-letter space followed by a capital letter is an eye-catching configuration that might induce pausing regardless of its role in our orthographic code; that even subjects who were treating the copying task as somehow only marginally a linguistic task might pause there simply because the signal is so salient. The right-hand column of Table 2 is offered as a counter to this hypothetical objection. Typed lines come to an end for non-linguistic reasons. The width of the paper obliges the typist, and the reader, to stop the rightwards traverse and return to the left margin. Notwithstanding this natural break which line ends present, subjects took them as opportunities for pauses only 27% of the time. They were clearly far more willing to copy on beyond the end of a typed line than to do so at the end of a sentence. Having been dealt with here, pausing at full stops is henceforth ignored, i.e. only sentence-internal pauses are considered in the analysis that follows.

	Percentages of opportunities taken to pause	
	at full stops	at ends of typed lines
COHERENT 'SEMANTICS'	82	27
COHERENT 'PHONETICS'	86	27

Table 2. Pausing at ends of sentences and at ends of typed lines.

For each of the four passages, subjects' pause records were amalgamated to produce pause distributions showing how

many subjects paused after each word in that passage. Product-moment correlation coefficients were used to evaluate the extent to which the distribution of pauses across the successive inter-word gaps within a given sentence matched the pause distribution of the syntactically comparable sentence in the other COHERENT passage. As control comparisons, product-moment correlations were also computed between the pause distribution for each sentence

		COHERENT 'SEMANTICS'				JUMBLED 'SEMANTICS'				COHERENT 'PHONETICS'			
		S_1	S_2	S_3	S_4	S'_1	S'_2	S'_3	S'_4	P_1	P_2	P_3	P_4
COHERENT 'SEMANTICS'	S_1												
	S_2	.54											
	S_3												
	S_4												
JUMBLED 'SEMANTICS'	S'_1	.18											
	S'_2		-.12										
	S'_3			.15									
	S'_4				-.28								
COHERENT 'PHONETICS'	P_1				.46								
	P_2			.69**									
	P_3	.32	.56*										
	P_4	.83**	.53									.37	
JUMBLED 'PHONETICS'	P'_1								.11	-.34			
	P'_2						-.47				-.13		
	P'_3					-.08						.28	
	P'_4					.17							.41

Table 3. Matrix of correlations for pause distributions in Experiment 1. ** highly significant, $p<.01$; * significant, $p<.05$. The sentences differ in length. Thus size of coefficient needed for significance varies. Unprimed abbreviations denote real sentences. Primedabbreviations are for the 'sentence counterparts' in the JUMBLED passages. Subscripts indicate the position within a passage of a given sentence or 'sentence counterpart'. Sentence-final pauses are excluded from the analysis.

and the distribution of pauses within its random 'sentence counterpart' in the corresponding JUMBLED passage, and between the pairs of pause distributions of these random 'sentence counterparts' across passages. Furthermore, since each of the COHERENT passages contained two equally long (14-word) sentences, it was possible to compute correlations comparing syntactically different structures, for these particular sentences. Table 3 (previous page) presents all the correlation coefficients.

The four underlined coefficients are the only ones in Table 3 which represent pause distributions for syntactically comparable sentences. Three of these four coefficients are significantly above chance level - two of them highly so. The remaining sixteen coefficients in Table 3 are all ones which were computed as control comparisons and none of them rises above chance level.

Discussion

Table 1 shows that some property, or combination of properties, of the COHERENT passages made them easier to copy in larger chunks than the JUMBLED passages. Table 2 shows that subjects were very sensitive to the division of the COHERENT passages into sentences. And Table 3 provides good grounds for believing that the internal structure of sentences is the major determiner of where subjects stop writing and look back for more when they are copying. The argument is that (with one exception which will receive further discussion below, viz. S_4xP_1) the pause distribution of a sentence correlates strongly only with that sentence in the other passage which was designed to exhibit comparable constituent structure. It fails to correlate significantly with the pause distribution for a random stretch of words located in the corresponding position in the JUMBLED version of the passage (and pairs of these random 'sentence counterparts' do not yield significant intercorrelations). Furthermore, for the 14-word sentences, the correlation coefficients fail to reach significance when sentences with different syntactic structures are compared.

Figure 3 presents tree diagrams derived from the pause distributions of the sentences in the COHERENT passages. Pairs of sentences intended to have comparable structures arealigned and the pauses falling in corresponding gaps are summed to provide the data upon which the trees are based. The tree for a pair of sentences represents their shared structure to the extent that it emerges through experimental 'noise'. The non-terminal node labels have been added on a purely intuitive basis whenever these appear to be justified by the strings they dominate (given their particular sentence contexts). It is my contention that the mere possibility of straightforwardly supplying theoretical linguistic labels for 30 nodes out of 48 (ie 62.5%) is a strong indicator that linguistic constituency influenced

pausing. Incidental support for this position is gained by examining parts of the sentences that were not fully syntactically comparable across the two passages. These design blemishes are the underlined strings of words in Figure 3. The relevant observations are: in S_1 there is only one pause within the NP common$_1$knowledge, but four

Figure 3. Tree diagrams for the amalagamated pause distributions of the sentences in the Experiment 1 COHERENT passages. The relevant correlation coefficients from Table 3 are reproduced beside the trees. Parts of thesentences that are not fully syntactically comparable across the two passages are underlined. A few long words have been contracted in sentences S_3 and P_2. See Figure 2 for full versions of these two sentences.

pauses separate the adverb and the verb in the falsely matched string generally$_4$recognized of P_4; in S_2 ...$_0$majority$_4$[of$_0$cases] received a distinct bracketing from its false match in P_3, ...$_5$[of$_1$the$_1$time], with the preposition phrases being respected both times; the subject NP of the first clause of P_2, the consonants is separated by four pauses from its VP, are formed, but the linearly corresponding gap in S_3, between does and not, attracted no pauses (a difference that many theoretical linguists might have predicted); and within the relative clauses of S_4 and P_1 the individual bracketings almost approach being theoretically reasonable,

...[may$_0$lead]]$_3$[to$_0$changes]$_3$[in$_1$meaning]
versus
...[has$_0$made]$_4$vowels]$_5$[difficult$_4$[to$_2$transcribe]].

These observations suggest that the text-copying method is considerably more sensitive to syntactic detail than had been anticipated. The mismatches underlined in Figure 3 are design blemishes only when the criteria for syntactic comparability exceed a certain level of delicacy. When the passages were being prepared no great sensitivity to detail was expected and it was assumed that one 7-word relative clause was as good as any other 7-word relative clause, one 6-word adverbial phrase a suitable match for any other 6-word adverbial phrase, etc. It is very likely that the four crucial correlation coefficients in Table 3 would have been enhanced in size if the pairs of sentences had been more closely matched. The S_4xP_1 correlation, which at 0.46 just fails to reach significance ($.10>p>.05$), could well have been raised to significance if the two relative clauses had had the same internal structure. In Experiment 2, reported below, closer structural parallels were striven for.

An important caution is evident from inspection of Figure 3: the text-copying method is clearly rather insensitive to syntactic detail at the beginnings of sentences; at least this appears to be true for highly literate subjects transferring, on average, four words of text at a time. Even though It at the start of sentences S_1 and P_4 is a major constituent, no subjects were content to take in only that one short word and write it down before proceeding to deal with more of the passage. After the first few words of a sentence, however, it appears that the fact that different subjects copied sequences of different lengths led to asynchronies which ensured a fair distribution of pauses among the remaining constituent breaks.

Word-length

Content words tend to be longer than grammatical function words and, in English, phrases generally start with a grammatical function word and end with a content word (eg

a noun phrase such as an armadillo is opened by a short function word and closed by a long content word). Conceivably the patterns exhibited in Figure 3 are simply a reflection of word-length: subjects take in one or more short (function) words and then, still having short-term memory space to spare, they take in a following long (content) word: this fills the available capacity and the sequence is then copied out before they pause and return to the passage for another chunk of text.[7] If this were so, one might either dismiss the results as artefactual or, alternatively, develop a speculation to the effect that the relationship between word-length and the function-word/content-word distinction was a natural adaptation to the ecological pressures of linguistic processing. In fact, however, the JUMBLED texts provide evidence that the function/content distinction is efficacious over and above its connection with word length. Of course, the differences in numbers of pauses between COHERENT and JUMBLED texts (Table 1) already indicate that something else in addition to mere word-length is implicated.

Figure 4 (opposite) shows the median numbers of pauses after words of different lengths in the two JUMBLED passages. (Medians are used because some of the distributions are skewed. The oddity of providing medians for the two 5-syllable words and the lone 6-syllable word is recognized.) The open circles depict the pattern for all 122 words (the two passage final words are omitted from the analysis). Numbers of pauses clearly do not simply increase in proportion to the number of syllables in a word. Rather, there is an enormous leap between 1-syllable words and 2-syllable words. The relationship could be idealized by saying that monosyllabic words attract small numbers of following pauses (Median = 2.24), but for words of more than one syllable the number of pauses fluctuates on either side of a high plateau. The horizontal line at 5.81, representing the median for 47 of the 48 polysyllabic words is an indication of where this plateau might be. (The preposition between, the only polysyllabic function word that happened to appear in these passages was omitted from the calculation. Its inclusion would scarcely alter the figure of 5.81 for the median.)

The 74 monosyllabic words comprised 16 content words (nouns, full verbs, adjectives and adverbs) and 58 function words. Their two separate medians (3.50 and 2.00, respectively) are given as xs in Figure 4. A median-test comparison of the two types shows that monosyllabic content words attracted a highly significantly greater number of pauses than monosyllabic function words (X^2_c=6.82, $p<.01$, df=1). Thus 1-syllable content words reach up towards the plateau for polysyllabic words. They do not quite reach it, however. By the median test, they attract significantly fewer pauses than do the 47 polysyllabic content words (X^2_c=5.04, $p<.05$, df=1).

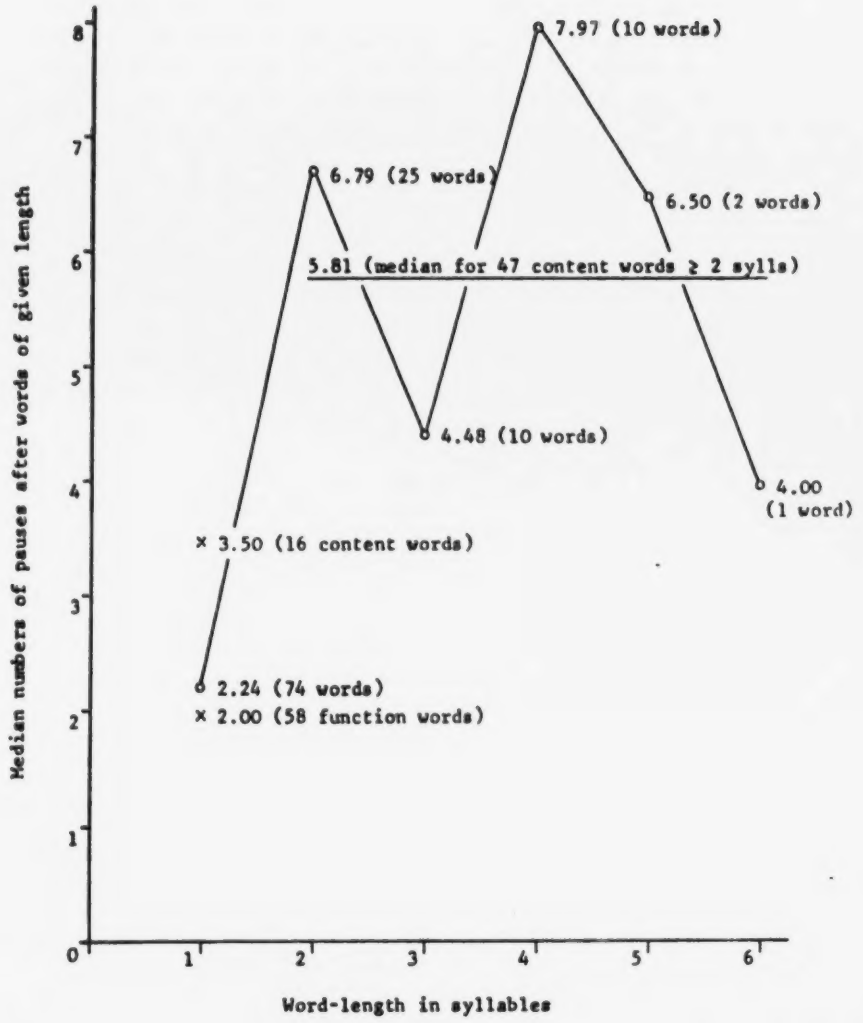


Figure 4. Median numbers of pauses after words of different lengths in the JUMBLED passages of Experiment 1. Passage-final pauses are excluded. Open circles are data points for all words. The xs show content words and function words separately, for monosyllables. For each data point the number of contributing words is indicated in brackets.

Any given random ordering is bound to yield some sequences that look like valid constructions. See, for example, the apparent VP can misclassify a vowel in the third line of the JUMBLED 'PHONETICS' passage (figure 2). To isolate word effects from constituent structure effects it would have been better to have given the JUMBLED passages in a different random order for every subject. Practical constraints dictated that only a single randomized version of each passage be prepared. Nonetheless it appears

incontestable that although word-length does influence pausing - manifested in the difference between monosyllabic and polysyllabic content words - the syntactically relevant distinction between content words and function words is also an important (and perhaps greater) factor.

Experiment 2

In view of the text-copying method's sensitivity to the details of constituent structure, suggested by the results of Experiment 1, it was decided to employ it with sentences exhibiting closer structural matches across passages. The second experiment used only COHERENT passages, because they had already been clearly demonstrated to evoke different behaviour from that seen with the JUMBLED passages.

An attempt was also to be made to use the text-copying method to elucidate the constituency of two particular construction types, regarding the analysis of which there is contention among theoretical linguists. Firstly, prepositional verbs (e g allow for). Does the preposition form "a semantic and syntactic unit with the verb ...." similarly to phrasal verbs such as find out, as claimed by Quirk, Greenbaum, Leech & Svartvik (1972:812) or is the preposition in closer construction with a following noun phrase, as it would be in an ordinary verb-followed-by-preposition-phrase-adjunct construction (V-PP) such as: speak[for a period of several seconds]? The question is whether prepositional verbs, such as allow for are appropriately bracketed [allow for] or allow [for NP].[8] Secondly, it seemed worth trying to find evidence for differences between the constituent structures of restrictive and non-restrictive relative clauses.[9]

A final aim of Experiment 2 was to investigate whether structurally ambiguous sentences would yield distinct pause distributions in contexts designed to favour different interpretations of them.

Subjects

Ninety-four senior undergraduates and beginning postgraduate Linguistics students participated in the experiment in fulfilment of a Psycholinguistics course requirement. None of them had participated in Experiment 1. As before, the experiment was conducted prior to class discussion of psycholinguistic studies of constituent structure. Half of the group, 47, acted as subjects and the remaining 47 were observers.

Materials

The eight typed 89-word passages used are reproduced in Figure 5 (overleaf). Each passage contained 8 sentences. The passages were meant to be read as coherent wholes, but

each sentence occupied a single line on its own, to avoid any confounding effects from typed lines coming to an end in mid-sentence. Double-line spacing was employed to facilitate reading. Capital letters and full stops were the only punctuation used.

The passages were designed as four pairs, 1 & 2, 3 & 4, 5 & 6, 7 & 8. Within the pairs there are only two differences: (a) the order in which the fifth and sixth sentences appear and (b) the final sentence in one member of the pair was meant to be taken as containing a restrictive relative clause (The man who ...) and in the other it was a sentence intended to be taken as containing a non-restrictive relative clause (Tom Man who ...). Within a given pair of passages the first seven sentences are identical.

The first four sentences are structurally parallel across all four pairs of passages. The fifth, sixth and seventh sentences of each passage were all 13 words in length and they were drawn from a set of: 2 sentences containing ordinary V-PP sequences (VPP_a speak for a period of several seconds, VPP_b stop at the limits of their knowledge); 3 sentences containing prepositional verbs (PrV_a contend with, PrV_b adjust to, PrV_c allow for); 3 sentences containing phrasal verbs (PhV_a make out, PhV_b find out, PhV_c put across); and two sentences intended to be ambiguous, one of them (Amb_a ... draw up the body of the tongue = 'raise between a phrasal verb analysis (in passages 1 and 2) and an ordinary V-PP analysis (in passages 3 and 4), and the other (Amb_b ... enlarge on a type of bromium paper = 'make enlargements on...' or 'say more about...') between a straightforward V-PP analysis (in passages 5 and 6) and a prepositional verb analysis (in passages 7 and 8). The manner in which these 10 sentences were partly counterbalanced across the fifth, sixth and seventh positions of the 8 passages is shown in Table 4.

	Passage Number							
Position	1	2	3	4	5	6	7	8
fifth	Amb_a	VPP_a	Amb_a	PhV_a	Amb_b	PrV_c	Amb_b	VPP_b
sixth	VPP_a	Amb_a	PhV_a	Amb_a	PrV_c	Amb_b	VPP_b	Amb_b
seventh	PrV_a	PrV_a	PrV_b	PrV_b	PhV_b	PhV_b	PhV_c	PhV_c

Table 4 Partial counterbalancing of structures presented as the fifth, sixth and seventh sentences of the 8 passages in Experiment 2. Abbreviations are explained in the text.

1. Traditional British phoneticians study the postures of the speech organs.
One needs to alter the shapes of the articulators in a skilled way to make speech sounds.
The lips must be moved.
We have to touch various positions on the palate.
For velar sounds we have to draw up the body of the tongue.
In normal situations people need to speak for a period of several seconds.
When making observations scientists have to contend with the speed of the performance.
The man who pioneered these studies died last year.

2. Traditional British phoneticians study the postures of the speech organs.
One needs to alter the shapes of the articulators in a skilled way to make speech sounds.
The lips must be moved.
We have to touch various positions on the palate.
In normal situations people need to speak for a period of several seconds.
For velar sounds we have to draw up the body of the tongue.
When making observations scientists have to contend with the speed of the performance.
Tom Man who pioneered these studies died last year.

3. X-ray cine photography reveals the motions of the speech organs.
One needs to mark the surface of an articulator with a barium pencil to make it visible.
The lips must be marked.
We have to make some lines on the palate.
For velar sounds we have to draw up the body of the tongue.
On exposed plates investigators need to make out the shape of each organ.
Before starting measurements subjects have to adjust to the taste of the barium.
The man who pioneered these studies died last year.

4. X-ray cine photography reveals the motions of the speech organs.
One needs to mark the surface of an articulator with a barium pencil to make it visible.
The lips must be marked.
We have to make some lines on the palate.
On exposed plates investigators need to make out the shape of each organ.
For velar sounds we have to draw up the body of the tongue.
Before starting measurements subjects have to adjust to the taste of the barium.
Tom Man who pioneered these studies died last year.

5. Detailed X-ray photographs reveal the nature of the vocal tract.
One needs to increase the size of an image with a convex lens to understand it easily.
The glottis must be widened.
We have to expand the gaps between the articulators.
For various reasons we have to enlarge on a type of bromium paper.
In making measurements investigators need to allow for the blurring of the edges.
By tracing contours phoneticians try to find out the volumes of the cavities.
The man who pioneered these studies died last year.

6. Detailed X-ray photographs reveal the nature of the vocal tract.
One needs to increase the size of an image with a convex lens to understand it easily.
The glottis must be widened.
We have to expand the gaps between the articulators.
In making measurements investigators need to allow for the blurring of the edges.
For various reasons we have to enlarge on a type of bromium paper.
By tracing contours phoneticians try to find out the volumes of the cavities.
Tom Man who pioneered these studies died last year.

7. Many scientific techniques confront the student of the human brain.
One needs to explain the nuances in the methods with great care to make them clear.
The materials must be described.
We have to specify a lot of the details.
For various reasons we have to enlarge on a type of bromium paper.
When studying psychology linguists like to stop at the limits of their knowledge.
After easy introductions teachers have to put across the complications of the subject.
Tom Man who pioneered these studies died last year.

8. Many scientific techniques confront the student of the human brain.
One needs to explain the nuances in the methods with great care to make them clear.
The materials must be described.
We have to specify a lot of the details.
When studying psychology linguists like to stop at the limits of their knowledge.
For various reasons we have to enlarge on a type of bromium paper.
After easy introductions teachers have to put across the complications of the subject.
The man who pioneered these studies died last year.

Figure 5 The eight passages used in Experiment 2. The format of each passage is shown exactly as it was presented to the subjects. The second sentence in passages 7 and 8 has only 16 words. It should have been 17 words long to match the second sentence in the other six passages.

Procedure

Testing took place in four separate sessions (with any given subject attending only one of the sessions) in a language laboratory. Each subject copied one of the eight passages. Assignment of subjects to passages was arbitrary and an attempt was made to ensure that each of the passages was copied by an equal number of subjects at each testing session. The positioning of subjects, observers and pieces of paper was exactly as described for Experiment 1.

After the group had assembled at a testing session, the subjects were withdrawn into a nearby room and subjects and observers were then given separate oral instructions.[11] (This made it easier to answer observers' requests for clarification without alerting subjects to the true nature of the task, than had been the case with the tape recorded instructions in Experiment 1.) Apart from obvious changes relating to the fact that each subject now had only one passage to copy, the instructions were the same for Experiment 1. As soon as they had been briefed, the subjects returned to the language laboratory and testing began.

Results

Table 5 gives the means and standard deviations of the number of pauses per subject for the four pairs of passages.

	Passages			
	1&2	3&4	5&6	7&8
mean	26.6	23.5	23.6	26.6
sd	7.6	7.3	11.9	10.8

Table 5 Means and standard deviations of pauses per subject for the four pairs of passages in Experiment 2. Passage-final pauses are not counted here.

An analysis of variance failed to reveal differences between these four means ($F<1$, i e, there was more variance amongst subjects copying passages in a given pair than there was from one pair to another pair). This implies that the passages were comparable in level of difficulty, a desirable outcome to have achieved. (Inspection had shown no large differences between the means within each pair of passages.) The overall average length of the sequences that subjects were carrying across before a pause was 3.40 words, which is

almost identical with the corresponding mean for the COHERENT 'PHONETICS' passage of Experiment 1 (3.41 words). And, also similarly to the 'PHONETICS' passages, a sprinkling of pauses within words was again recorded. These were treated in the same way as they had been in the analysis of Experiment 1.

As had been found in the first experiment, subjects were strongly inclined to pause at the ends of sentences. Altogether they were recorded as having paused 269 times at full stops (excluding the last one in each passage, where pausing was unavoidable). Out of 329 occasions when they could have done so (47 subjects x 7 sentences), this represents opportunities being taken 82% of the time.

In an experiment such as this one, observer accuracy is obviously a matter for concern and Table 6 presents one check on the quality of observers' records. Some of the sentences were quite long and subjects' written copies of these often extended over more than one line. Was there a tendency for observers to confound instances of subjects lifting their pens to return to the left hand margin with pauses for looking back at the original text? Table 6 classifies every gap between words, except for gaps containing full stops. It turns out that the majority of recorded pauses did not coincide with the ends of written lines. There is essentially no dependency between the two dichotomies (Phi coefficient, r_ϕ=-.098). This potential distraction, at least, was not confounding the observers.

	Pause recorded	Pause not recorded
Subject ends a line of writing	115	190
Elsewhere	768	2643

Table 6 Relationship between the recording of pauses and whether or not subjects came to the end of a line of writing, in Experiment 2. Sentence-final pauses are excluded here, since subjects often followed the original passages in starting each sentence on a new line. For this table, n = 46 because one subject's written copy was mislaid.

Turning now to individual sentences and sentence types, the results will be presented in the following order: (1) the ambiguous sentences (2) the sentences containing

relative clauses (3) the first four sentences of the passages (4) the unambiguous sentences containing prepositional verbs, phrasal verbs and verbs followed by ordinary preposition phrases. The most interesting of these results are those dealt with in (4). The results covered in (3) form a baseline of comparison for (4). On the other hand, (1) and (2) are essentially reports on failures. They are included for the sake of completeness.

(1) The ambiguous sentences

The sequence draw up in passages 1 and 2, which were intended to provide a context favouring a phrasal verb analysis, attracted the pattern of pauses $draw_2up_7the$, i.e. draw up formed a constituent as expected. However, in passages 3 and 4, which were supposed to favour a V-PP analysis, it again yielded a pattern that tended towards the phrasal verb type: $draw_3up_5the$. Enlarge on, the other ambiguous sequence, gave the pattern $enlarge_3on_3a$ in passages 5 and 6, where it had been hoped that the context would favour V-PP. This pattern of course fails to connect on more closely with either of its neighbours. And in passages 7 and 8, which were supposed to provide pressure for a prepositional verb interpretation, the pattern was $enlarge_4on_0a$, which suggests a V-PP analysis. The results recounted in (4), below, establish that the latter is probably the correct analysis for verb phrases having prepositional verbs in them, but that is small comfort.

Inadequately constructed test passages vitiate the comparisons being attempted here. Post hoc, the passages were presented to 49 Linguistics students who had not been copiers in either of the experiments. The ambiguous sentence in each passage was underlined. They were asked to read through the passage until they reached the underlined sentence and then write down a paraphrase that represented their "first reaction as to what the sentence means". Each of them read only one passage. Nine of the judges answered in ways that could not be relevantly interpreted, eg. by giving paraphrases that were themselves ambiguous on the crucial issue. Of the remaining forty, 23 judged the draw up sentence. Overwhelmingly, it was taken as containing a phrasal verb. Twenty judges paraphrased it this way regardless of which context it had appeared in, and only 3 - all of whom read it in the V-PP favouring contexts - offered V-PP paraphrases. For enlarge on, the preferred interpretation was V-PP. All 7 of the judges who read enlarge on in the V-PP favouring passages took it this way, but so did 4 of the 10 who had it in passages slanted towards the prepositional verb analysis. Obviously it would have been better to have had these sentences judged by the copiers themselves, immediately after the experiment. Individual judgements and pause patterns could then have been compared. And test passages providing more persuasive contextual pressure would have been desirable.

There is one small indication that the pause distributions for the ambiguous sentences might have been influenced by context: rank correlations[12] comparing the pause distributions of each of these sentences across its two types of context fail to reach significance (for ... draw up ...., $r'=0.46$, $p>.10$; for ... enlarge on ...., $r'=0.41$, $p>.10$). Compared to the rank correlations reported in Table 8 below, for very similar sentences, these non-significant coefficients would be surprisingly low, unless disambiguating contextual pressures were implicated.

(2) Relative clauses

The final sentence in each passage was either The man who pioneered these studies died last year (relative clause intended, in the contexts provided by the passages, to be restrictive) or Tom Man who pioneered these studies died last year (relative clause intended as non-restrictive). The empirical major constituent break for both sentences turned out not to be after the relative clause, nor before the relative clause, nor in any other reasonably expectable place, but halfway through the relative clause (14 out of 24 subjects paused after pioneered in the restrictive version and 14 out of 23 paused after pioneered in the non-restrictive version). This puzzling finding will receive further comment in the CONCLUDING DISCUSSION, below. What happened to the left of this strange break might nonetheless be considered to be of interest for the restrictive/non-restrictive distinction. The 24 subjects who copied the putative restrictive version produced the following pattern: $[\text{The}_2\text{man}_2\text{who}_2\text{pioneered}]_{14}$. The other 23 subjects gave: $[\text{Tom}_5\text{Man}]_7[\text{who}_5\text{pioneered}]_{14}$, for the intended non-restrictive version. It would not be impossible to argue that this difference is rational, with the cluster of 7 pauses occurring where the first comma would standardly be written to demarcate a non-restrictive relative clause. To do so would be opportunistic, however. Notice that the first three interword gaps of the 'restrictive' version contain only 6 pauses and that the corresponding part of the 'non-restrictive' one has 17 pauses. Perhaps the 'non-restrictive' version was somehow peculiar, eg. for having been typed without commas or because Tom Man is not a name that 'rings a bell' as a famous phonetician/radiologist, etc. Peculiarity might have led to 'double-take' pausing unrelated to the goals of the experiment.

These two relative-clause sentences yielded pause distributions that are overall very similar. The rank correlation coefficient, r', is 0.88 ($p<.02$). The fourth sentence in each passage was the same length as the relative-clause sentences, 9 words, but structurally different (see Figure 5). The amalgamated pause distribution for the 4 fourth-position sentences differs significantly from the combined distribution for the 2 relative-clause sentences ($X^2=16.36$, $df=7$, $p<.05$; and

inspection suggests that one slightly undersized 'expected' cell entry does not detract from the significance of the result). This difference tends to confirm the text-copying method's sensitivity to constituent structure, even if in the present experiment it has told us hardly anything about relative clauses.

(3) The first four sentences

The average rank correlation coefficients[13] between the pause distributions for each of the first four sentences, across pairs of passages, are displayed in Table 7. The second sentence in passages 7 and 8 was only 16 words long; it should have contained 17 words to match the second sentence in each of the other 6 passages (see Figure 5). The disparity arises after the preposition in the phrase with great care (cf. in a skilled way, with a barium pencil, with a convex lens). From the 11 subjects who copied this maverick sentence, the total number of pauses between with and great was 2. For purposes of analysis, this cluster of 2 pauses was treated as if it had been two single pauses on either side of an imaginary word, i.e. with$_1$imaginaryword$_1$great. There is no reason at all to think that this marginal manoeuvre had any noticeable effect on the general tenor of the results.

	Sentence position in passages			
Sentence position in passages	1st	2nd	3rd	4th
1st	.62**			
2nd		.67**		
3rd			.20	
4th				.21

Table 7. Average rank correlation coefficients, $\bar{r}'$, between the pause distributions, across pairs of passages, for each of the first four sentences in the Experiment 2 passages. ** highly significant, $p<.01$. Sentences differ in length. Thus size of coefficient needed for significance varies.

As Table 7 shows, there is a satisfyingly high leel of agreement amongst the pause distributions of the four sentences that appeared in initial position in each passage ($\bar{r}'=0.62, p<.01$); likewise for the set of four sentences in second position ($\bar{r}'=0.67$, $p<.01$). Tree diagrams, with non-

terminal nodes intuitively labelled, for these first two sets of sentences are presented in figures 6 and 7. The coefficients for the sentences in third position and in fourth position fail to reach significance, though they do begin to approach that level (.20>p>.10, in both cases). The small measure of agreement amongst the pause distributions for the sentences in third position is almost certainly attributable to the fact that these sentences were too short (5 words long) to attract a reasonable number of pauses from subjects as literate as advanced university students. Overall, these four sentences received only 22 sentence-internal pauses from the 47 subjects, an average of less than half a pause per subject per sentence. A tree diagram based on the amalgamated distribution for these five-word sentences, for what it is worth, was presented as an illustration in the RATIONALE section above (Figure 1).

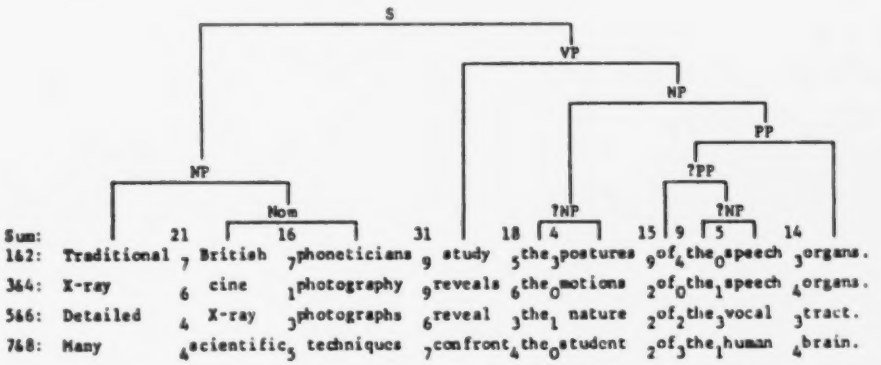

S
VP
NP
PP
NP
?PP
Nom
?NP
?NP
Sum: 21 16 31 18 4 15 9 5 14
162: Traditional 7 British 7 phoneticians 9 study 5 the 3 postures 9 of 4 the 0 speech 3 organs.
364: X-ray 6 cine 1 photography 9 reveals 6 the 0 motions 2 of 0 the 1 speech 4 organs.
566: Detailed 4 X-ray 3 photographs 6 reveal 3 the 1 nature 2 of 2 the 3 vocal 3 tract.
768: Many 4 scientific 5 techniques 7 confront 4 the 0 student 2 of 3 the 1 human 4 brain.

Figure 6. Tree diagram for the amalgamated pause distribution of the set of four sentences that appeared in initial position in the Experiment 2 passages. Numbers to the left identify the passages in which they were used.

It is difficult to see why the sentences in fourth position failed to elicit a significant level of agreement - though, at only 9 words long, they do tie with the relative-clause sentences in being the second shortest ones used. A disparity in the numbers of pauses falling between the object NP and the following PP in these sentences does seem explicable, however. Pausing in this gap was as follows: positions$_7$on, lines$_6$on, gaps$_4$between and lot$_1$of. Unlike the others, the sequence lot of is an idiomatic combination, as testified by the possibility of phonological cliticization represented in the non-standard spelling lotta. And I assume that idiomaticity is a route which can lead to reanalysis of constituent structure. Omitting the passage 7 and 8 lotta sentence from the calculation yields a

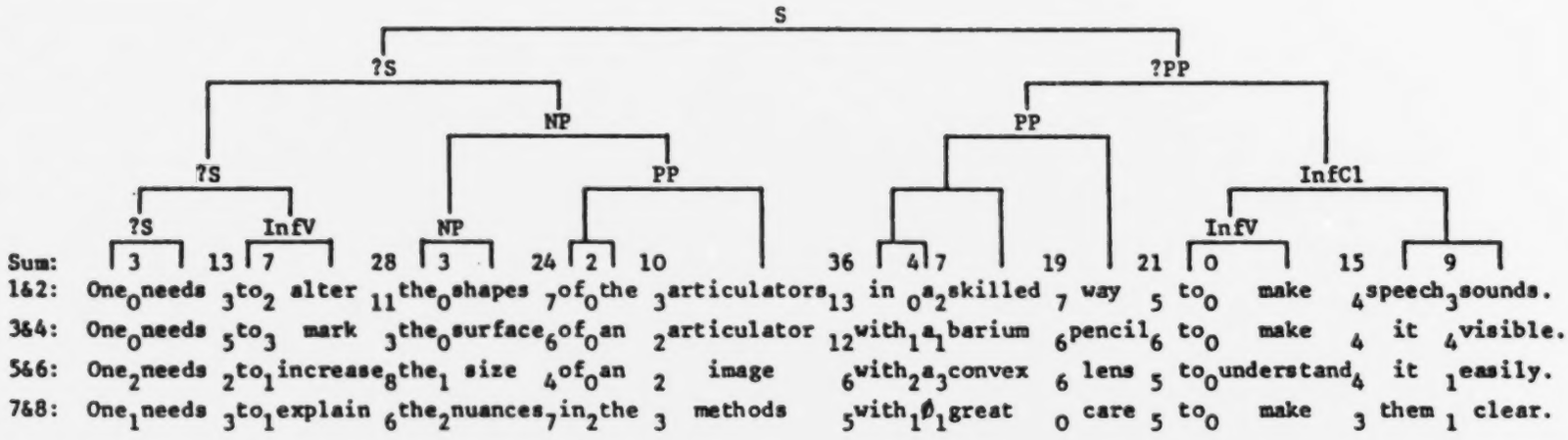

S
?S
?PP
NP
PP
?S
PP
InfCl
?S
InfV
NP
InfV
Sum: 3 13 7 28 3 24 2 10 36 4 7 19 21 0 15 9
1&2: One0needs 3to2 alter 11the0shapes 7of0the 3articulators13 in 0a2skilled 7 way 5 to0 make 4speech3sounds.
3&4: One0needs 5to3 mark 3the0surface6of0an 2articulator 12with1a1barium 6pencil6 to0 make 4 it 4visible.
5&6: One2needs 2to1increase8the1 size 4of0an 2 image 6with2a3convex 6 lens 5 to0understand4 it 1easily.
7&8: One1needs 3to1explain 6the2nuances7in2the 3 methods 5with1a1great 0 care 5 to0 make 3 them 1 clear.

Figure 7. Tree diagram for the amalgamated pause distribution of the set of four sentences that appeared in second position in the Experiment 2 passages. Numbers to the left identify the passages in which they were used.

higher average rank correlation for the other three fourth-position sentences; one which now definitely approaches significance ($\bar{r}'$=0.48, .10>p>.05). The empirical constituent structure of these sentences is postponed for examination in the CONCLUDING DISCUSSION below (see Figure 9).

(4) Unambiguous sentences with prepositional verbs, phrasal verbs and ordinary V-PP sequences

These sentences, which appeared in various orders and combinations (together with the ambiguous sentences) in fifth, sixth and seventh positions in the passages (see Table 4), produced the clearest and most encouraging results of Experiment 2. Table 8 presents the relevant rank correlations and X^2 comparisons between the three types of sentence. The coefficients on the main diagonal show that within the sets of sentences representing each of the three types there was a highly significant level of agreement. The other statistics show that the amalgamated pause distribution for the sentences with phrasal verbs was very clearly different from those for the other two types of sentence (see the X^2s in the bottom row of Table 8), while the amalgamated distribution for sentences with prepositional verbs was essentially indistinguishable from the one for sentences with ordinary V-PP sequences (a highly significant rank correlation and a completely non-significant X^2 value, .90>p>.80). Regarding the highly significant X^2s in the bottom row of Table 8, inspection reveals that by far the largest component of X^2 in each case arises out of the verbs of phrasal verbs being in close

Type of structure	Type of structure: VPP (2 sentences)	PrV (3 sentences)	PhV (3 sentences)
VPP	r'=.79**		
PrV	r'=.73** X^2=2.50ns	$\bar{r}'$=.72**	
PhV	X^2=20.95**	X^2=34.12**	$\bar{r}'$=.73**

Table 8. Correlations and contrasts within and between the pause distributions for the unambiguous Experiment 2 sentences containing verbs followed by ordinary preposition phrase adjuncts (VPP), prepositional verbs (PrV) and phrasal verbs (PhV). ** highly significant, p<.01.

construction with the particles that immediately follow them, while in both of the other types of sentence the verbs are relatively separated from the following prepositions. A sizeable secondary X^2 component, in each case, comes from the particles of phrasal verbs being relatively disjointed from the object NP, whereas the prepositions in the other two types are closely associated with the NP which follows. Nearly all of the other components of these two X^2s are negligible in size.[14] The highly significant differences between whole sentences are therefore almost entirely attributable to the difference in backeting focused on in the design of this part of the experiment, viz: [[V PARTICLE]NP], for phrasal verbs, as opposed to V[PREP NP]. Tree diagrams for the three sets of sentences discussed here appear in Figure 8 (opposite). As before, the node labels have been added on an intuitive basis.

Grosjean et al (1979:74 & 76) report the same structure as above for the phrasal verbs call up and bring out in their study of pause durations in reading aloud. However, their method also gives this phrasal-verb-type bracketing to a prepositional verb that they tested (1979: 71): ...[[wondered about] this] [extraordinary story]. They attribute this result to the fact that about this extraordinary story is longer than wondered: "... pauses will coincide with this break [i e the boundary immediately after the verb] only if the verb and the following NP (or PP) are of equal length..." (1979: 74). Comparison on this point suggests that text-copying may be a better technique for revealing linguistic constituent structure. The three sentences providing the data for the middle (prepositional verb) diagram of Figure 8 have six-word PPs following the verb, but in all three cases the break immediately after the verb is marked by more pauses than fell at any of the breaks within the PP.

CONCLUDING DISCUSSION

Particular constructions

While there certainly are aspects of Figures 3, 6, 7, 8 and (see below) 9 that do not harmonize tidily with any distillation of theoretical linguists' descriptions of comparable sentences, there are numerous points of agreement as well. The large number of nodes in the tree diagrams that it has been straightforwardly possible to label support this claim. Sentence intial adverbials, sentence subjects and preposition phrases provide further support, though there are important qualifications to be mentioned in connection with the latter two.

In the two experiments, subjects saw a total of 13 sentences with initial adverbials (S_2, P_2 and P_3 in Experiment 1, and the Experiment 2 sentences in 5th, 6th and 7th positions; each of the ambiguous sentences being treated

here as a single sentence with its pause distribution being the sum across its different contexts). These adverbials were of various kinds: preposition phrases and finite and non-finite clauses. In 11 of the 13 sentences the pause distributions yielded bracketings in which the adverbial was a constituent, and furthermore an immediate constituent of

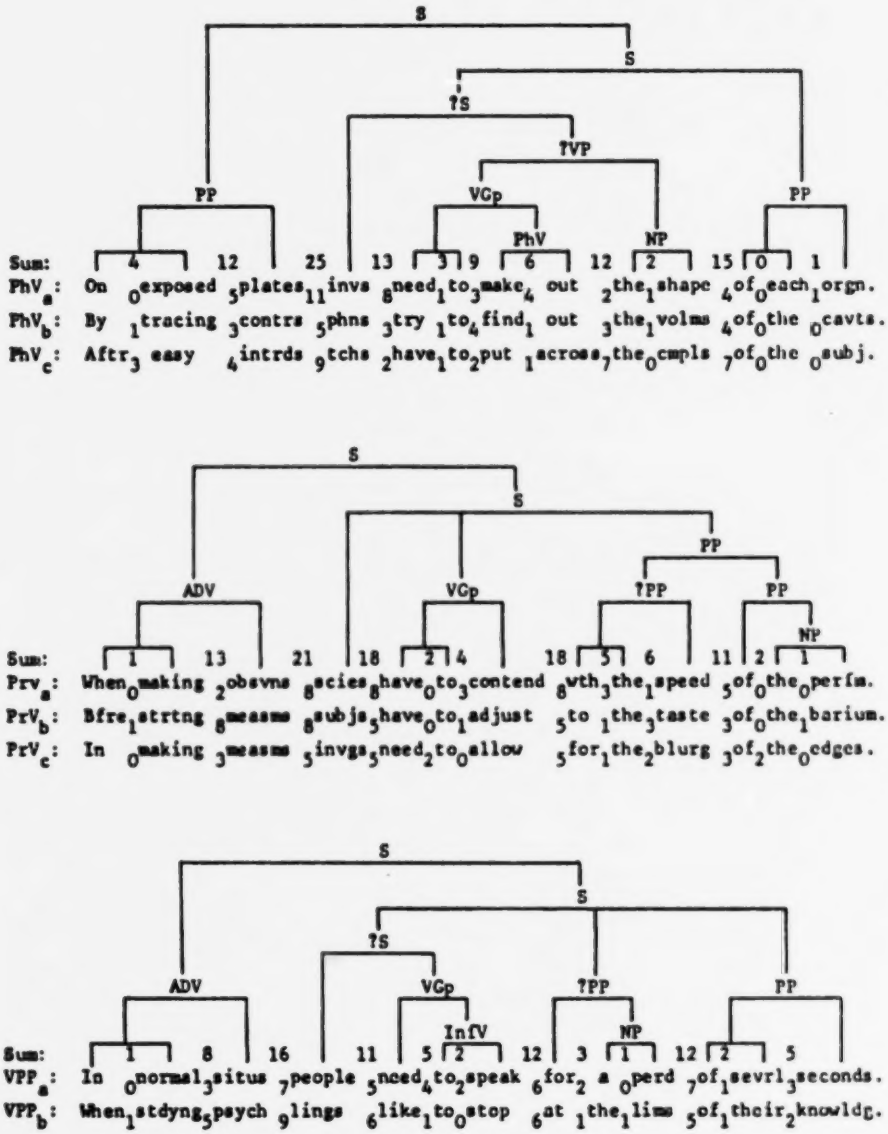

S
S
?S
?VP
PP
VGp
PP
PhV
NP
Sum: 4 12 25 13 3 9 6 12 2 15 0 1
PhVa: On 0exposed 5plates 11invs 8need 1to 3make 4 out 2the 1shape 4of 0each 1orgn.
PhVb: By 1tracing 3contrs 5phns 3try 1to 4find 1 out 3the 1volms 4of 0the 0cavts.
PhVc: Aftr 3 easy 4intrds 9tchs 2have 1to 2put 1across 7the 0cmpls 7of 0the 0subj.
S
S
PP
ADV
VGp
?PP
PP
NP
Sum: 1 13 21 18 2 4 18 5 6 11 2 1
Prva: When 0making 2obsvns 8scies 8have 0to 3contend 8vth 3the 1speed 5of 0the 0perfm.
PrVb: Bfre 1strtng 8measms 8subjs 5have 0to 1adjust 5to 1the 3taste 3of 0the 1barium.
PrVc: In 0making 3measms 5invgs 5need 2to 0allow 5for 1the 2blurg 3of 2the 0edges.
S
S
?S
ADV
VGp
?PP
PP
InfV
NP
Sum: 1 8 16 11 5 2 12 3 1 12 2 5
VPPa: In 0normal 3situs 7people 5need 4to 2speak 6for 2 a 0perd 7of 1sevrl 3seconds.
VPPb: When 1stdyng 5psych 9lings 6like 1to 0stop 6at 1the 1lims 5of 1their 2knowldg.

Figure 8. Tree diagrams for the Experiment 2 amalgamated pause distributions of the unambiguous sentences with phrasal verbs (top), prepositional verbs (centre) and verb followed by ordinary preposition phrase (lower diagram) structures. Some long words have been contracted. See Figure 5 for full versions of the sentences.

the whole sentence. For 7 of these 11 positive cases the correspondence with theoretical linguists' accounts is even closer, the emirical structure being $_S$[ADV[S]], i e the break after ADV is the major constituent break of the sentence.

It is a commonplace of introductory syntax classes that some students balk at accepting that ordinary transitive English sentences split naturally into two parts: a subject noun phrase and a predicate phrase (VP). And distributional arguments for this constituency woo them only reluctantly away from such alternative analyses as [[Subject+verb]Object]. This latter, unorthodox, bracketing has sometimes received a modicum of support from other empirical approaches to constituent structure, e g Martin (1970), Levelt (1970:121) and Grosjean et al (1979: 72). However, my pause data offer good evidence in favour of the distributionally justified analysis, except when the sentence subject is a pronoun. Ignoring clauses with pronominal subjects, there were in the two experiments 27 clause subjects, which presented the copiers with a total of 245 occasions on which they could pause immediately after copying a subject NP. In 149 instances they did opt to pause. That is to say that the constituent break after a full NP subject was respected 61% of the time, which approaches the figures of 82% and 86% of opportunities taken for pausing at the ends of sentences, as reported in earlier sections of this paper.

For the 17 clauses with pronominal subjects (it, this, one, we and the relative pronouns who and which) the result is quite different. This set of clauses also provided the copiers with a total of 245 opportunities for pausing immediately after the subject, yet they did so only 23 times, 9%. This behaviour in relation to pronominal subjects may be an artefact of the experimental method used, given that pronouns are short function words and that in 10 cases out of the 17 they were also the first word of a sentence (which means that they appeared in the position where the text-copying method shows least sensitivity). However, the number of languages in which subject pronouns are optional or are clitics on the verb suggests that it would be imprudent to dismiss out of hand this possible pointer to pronominal subjects being similarly different from full NP subjects in English too.

Altogether there were 53 preposition phrases in the test passages of the two experiments. (Here the preposition of a prepositional verb is counted as starting a preposition phrase, in line with the empirical findings reported above. Phrasal verb particles are, of course, not treated as prepositions. Again, each of the ambiguous sentences was regarded as a single sentence, notwithstanding its different contexts.) These preposition phrases emerged clearly as unified constituents: for 43 of the 53 (i e 81%)

there were fewer pauses at any point within the phrase than in either of the gaps flanking it.

An interesting comparison is possible here with some data published by Hawkins (1971) on hesitation pauses in a sample of children's spontaneous narrative speech. Of the 53 preposition phrases in the present test material, there were 39 that contained a determiner (e g between the articulators, with a convex lens, as opposed to e g of bromium paper, on exposed plates) and all but one of these 39 appeared in either final position or mediallyin their sentences. For the 38 non-initial preposition phrases containing determiners the overall amalgamated pause pattern was: $_{201}$PREP$_{46}$DET$_{53}$N or ADJ. Summing 201, 46 and 53 gives a total of 300 pauses adjacent to the prepositions and determiners. The pause pattern expressed in percentages of 300 is then $_{67\%}$PREP$_{15.3\%}$DET$_{17.7\%}$N or ADJ, and this is remarkably close to the corresponding figures derived from the last line of Hawkins' Table 1 (1971:283): $_{69.5\%}$PREP$_{13.3\%}$DET$_{17.1\%}$N. Notice that both sets of data would then seem to license the curious bracketing pp[[PREP DET]...]. However, this probably gives too much weight to the small different between 46 and 53 pauses (for Hawkins' data the raw figures are 14 and 18 pauses, respectively). Furthermore, when the 38 non-initial determiner-containing preposition phrases are examined individually it is found that, although 20 of them attracted fewer pauses to the gap preceding the determiner than fell immediately after it (i e [[PREP DET]....]) and only 11 had the converse pattern of more pauses before the determiner than after it (i e [PREP[DET....]]), this difference fails to be statistically significant (two-tailed sign test corrected for continuity, p=0.19). It therefore seems safest to conclude that, while the present results amply confirm the reality of preposition phrases as constituents, they do not reliably map the internal structure of preposition phrases.

Though experimental 'noise' (discussed further at the end of this paper; see Final remarks) masks some of the details and makes for some untidiness, the points listed above about particular constructions suggest that there is a fair measure of agreement between the empirical pause distributions and expectations derived from theoretical linguistics. What light then do the empirical results throw on structures for which there is not yet a consensus among theoretical linguists? One hoped-for clarification failed to materialize: for reasons already discussed, the Experiment 2 results do not firmly demonstrate any differences between sentences with restrictive relative clauses and those with non-restrictive relative clauses (which is not to say that there are no differences; just that the materials employed here were not good enough for the task of finding whatever differences may exist). Note, however, that both of the Experiment 1 sentences containing relative clauses, S_4 and P_1, yield the same pattern - and

one that is compatible with a popular analysis of restrictive relative clause structures - viz. $_{NP}$[[NP]WH....] (see Figure 3).

More positively, Experiment 2 gave an unequivocal verdict on the constituency of the prepositional verbs tested. Their structure is no different from that found in ordinary V-PP sequences, and it is sharply different from transitive phrasal verb constructions. No clear generalizations about the relationship of preposition phrases to other constituents emerge in the present results. However, it is interesting that the by-phrases in the two passive sentences tested (S_2 and P_3: see Figure 3) both turn up as 'daughters' of VP. Also, preposition phrases which act as nominal modifiers do several times, though not always, appear in close construction with the modified NP, e g S_1 and P_4 in Figure 3, and the tree diagram in Figure 6.

When a sentence contained one or more auxiliary verbs, a constituent comprising just the auxiliarie(s) and main verb generally came to light. This may be taken as support for the Hallidayan verbal group, and such constituents have accordingly been labelled VGp in Figures 3 and 8. See Halliday (1976) and Huddleston (1976:81 & ch14) for some relevant theoretical linguistic discussion of this matter. Martin (1970) found a strong predilection for this grouping and Grosjean et al (1979: 78) apparently observed a similar tendency too. An additional observation on constituency in which verbs participate concerns the infinitival marker to. The passages for Experiment 2 contained 22 instances of this to (as before, each of the ambiguous sentences is treated as a single sentence). In 14 of these cases to turned out empirically to be in close construction with what followed it, usually giving a constituent [to V], and in 6 cases to joined on to what preceded, usually forming a constituent [V to]. The remaining two examples had the empirical constituency [...V to V]. Although the number of pauses involved is small and the difference between the number of pauses before and after to was often slight, the preponderance in favour of the [to V (....)] bracketing is statistically significant (two-tailed Wilcoxon T-test, $p<.05$). Six of the seven instances in which need(s) was the first verb conformed to this pattern.

Comprehension in text-copying

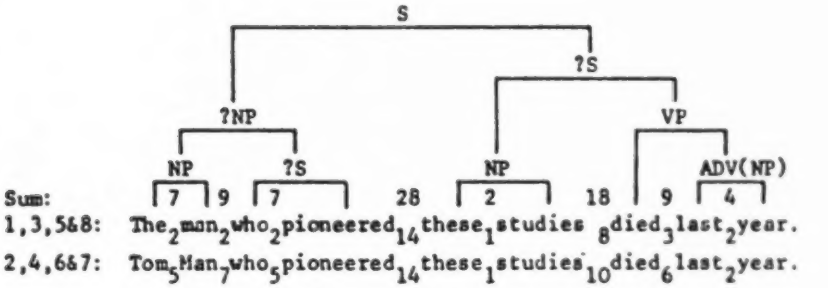
Immediately after the end of Experiment 2, the subjects were asked to write down any general comments they had and also to give a subjective impression of how well they had understood their test passage: "Did you understand what you were writing while you were doing it?" Six offered general remarks but neglected to comment on their understanding of the passage. That leaves 41 who did make notes relevant to this question. Only 13 of these 41 answers were negative, but even so the majority of negative answers were tempered

("I was not really understanding it as I wrote", "I didn't feel I was really taking it in", "I didn't understand what I was copying. I knew vaguely that it was about phonetics", etc). Only 2 of the 13 negative responses lacked this sort of qualification and they both said that they "would not have been able to answer a comprehension test" and one of them added "because I was writing quickly, without bothering to try and remember". Since memory for the purpose of answering a comprehension test would be something over and above mere comprehension at the time of copying, it is open to us to believe that even these two subjects perhaps understood a little of what they were copying. Varying degrees of positive comprehension were claimed by the remaining 28 subjects (i e by 68% of the 41 who answered the question). According to their own reports, therefore, many - perhaps nearly all of the copiers (remember the qualifications embedded in the majority of negative responses) - showed at least some understanding of the passages, which is to say that the task was apparently one that, possibly to varying extents, did engage high level linguistic analyzing mechanisms.

On the other hand, only 5 of the 41 subjects made unqualified affirmative claims regarding their comprehension ("I understood what I was writing", "The material was easy to understand while I was copying", "I understood the passage", etc). If it is accepted that the Experiment 2 passages were not so contrived and to be nearly incomprehensible, it is worth asking why so many subjects found cause to dilute the strength of their affirmations ("The last sentence was the only one I properly understood. The others I understood fragments of", "In most cases I think I understood", "I didn't get the whole thing but, yes, a part of it", "Vaguely understood the passage", etc). For 6 subjects the issue is fairly clear: they reported comprehension of the individual sentences but failure to integrate the sentences into a coherent whole ("Had some idea of what I was writing about, but limited to the sentence rather than the overall structure of the whole passage", "I understood the text whilst copying, but [....] didn't relate the content of the various sentences", etc).

These 6, plus 5 unqualified affirmatives, still leave unaccounted for 30 out of 41 (or 28 out of 41 if the two who doubted that they would succeed in a comprehension test are presumed to have understood nothing) who somehow felt that they had not <u>properly</u> grasped the meaning of what they copied, notwithstanding the fact that they were intelligent, highly-literate, experienced readers. Why? Perhaps because their language analyzing mechanisms were only partly engaged. If so, it behoves us to ask which parts were in play and which were idling. The tree diagrams in Figure 9 suggest a possible answer to this question. They represent the amalgamated pause distribution for the Experiment 2 relative-clause sentences and that for the set of four

sentences which appeard in fourth position in the Experiment 2 passages. Amalgamating the pause distributions for the fourth-position sentences and representing them by means of a single tree diagram is a questionable move in view of the low correlations amongst these four sentences ($\bar{r}'$=0.21, as shown in Table 7), but the aspects of this tree commented upon below are not seriously affected by disparities within this set of four distributions.



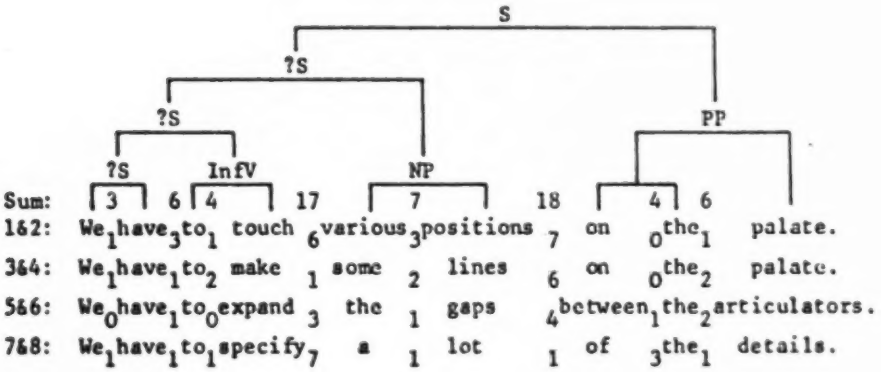

S
?S
?S
PP
?S
InfV
NP
Sum: 3 6 4 17 7 18 4 6
1&2: We1have3to1 touch 6various3positions 7 on 0the1 palate.
3&4: We1have1to2 make 1 some 2 lines 6 on 0the2 palate.
5&6: We0have1to0expand 3 the 1 gaps 4between1the2articulators.
7&8: We1have1to1specify7 a 1 lot 1 of 3the1 details.

Figure 9. Tree diagrams for the amalgamated pause distributions of the Experiment 2 sentences containing relative clauses and for the set of four sentences which appeared in fourth position in the passages. Numbers to the left identify the passages in which they were used.

The tree diagrams of Figure 9 suggest that some subjects were 'jumping to confusions' about constituency at the earliest possible opportunity. The man who pioneered could be a subject NP (it would be in a sentence such as The man who pioneered had to be self-reliant). Someone reading the passage for sense would discover from the sequel, these studies died last year, that there was more to the subject

NP: that these studies was the object of the verb in its relative clause. But someone who merely had to transfer a convenient chunk of text into memory for the purpose of making a copy might well be satisfied to grab the first string that could possibly form a constituent, without studying the remainder of the sentence to check that it was in fact a constituent of the sentence in hand. This is perhaps what was being alluded to by the following two answers to the post-test question about comprehension: "I did understand the sentences, but only after I'd written them in their entirety" and "I waited and checked what I was copying because I wanted to make some sense out of it". Obiously comprehension will be fragmentary if one takes The man who pioneered to be a subject NP that has somehow lost its predicate and then reads the string these studies died last year as a separate sentence in its own right.

Likewise in the lower diagram of Figure 9, We have to touch is a possible constituent (in a different context it could be a full clause) and so is We have to touch various positions, and perhaps even the fragment We have. But they - and the corresponding fragments of the other three sentences - are strings whose adoption as constituents would confound attempts to extract the full meaning of the particular sentences containing them here. Inspection of other tree diagrams presented earlier in this paper reveals further instances of this phenomenon, for instance the 'false' S constituents in the left half of Figure 7 and the 'false' adverbial Before starting in Figure 8.

For text-copying as an empirical technique to reveal the constituent structure of sentences, these observations have nuisance value (which could presumably be abated by careful preliminary editing of the test passages and by telling copiers that they will be asked comprehension questions afterwards). However, they are more interesting when viewed as clues to the nature of human on-line linguistic analysis. Kimball (1973) put forward several hypotheses as to how language users do surface structure parsing. One of these, his Principle 5, is clearly applicable here: "a phrase is closed as soon as possible, i e, unless the next node parsed is an immediate constituent of that phrase". The present results suggest that Kimball's Principle 5 can be decomposed into two subprocesses: 'close a phrase as soon as possible' and 'check that the next node parsed is not an immediate constituent of the last phrase thus closed'. The decomposability has come to light because text-copying does not demand full comprehension: easily memorizable constituents can be extracted from a text even by a subject ignoring the unless-clause of Principle 5. This interpretation is obviously also compatible with Frazier & Fodor's (1978) two-stage parsing model. I surmise that the kind of minimal comprehension required for efficient text-copying provides a laboratory example of the embarrassment

of the tired reader who comes to the foot of a page and then has to return to the top to re-read the page for a proper understanding of it.

There is an additional datum in the comprehension results which, though it certainly does not reach statistical significance here, would be worth investigating further. This is that 3 of the 5 copiers who made unhedged claims to the effect that they had understood what they were writing out all paused less often than the median number of pauses per copier and, by contrast, 8 of the 13 whose statements about their comprehension were negative had individual pause totals greater than the median.

Final remarks

For the 70 copiers in the two experiments there were 70 different observers. The observers were furthermore untrained, their task (but not the hypotheses) having been explained to them only minutes before they began making observations. This arrangement had the desirable consequence of rendering it extremely unlikely that the constituents which emerged in the aggregated pause distributions were the products of a wishful experimenter consciously or unconsciously imposing a favourite theoretical linguistic analysis in the process of deciding what to record as glancing-back-to-the-original-passage pauses. But the pooled observations of 70 inexperienced observers might easily have been uninterpretably chaotic. Undoubtedly, variations between observers and the fact that they had only a single quick chance to observe each pause must have generated considerable 'noise' in the data. The extent to which intuitively reasonable structures nonetheless penetrated the noise is all the more impressive. It testifies both to the salience of syntactic constituents as coding units in linguistic processing and to the robustness of the test-copying method. It should be noted that this noise is not indissociably linked to the method: cleaner data could doubtlessly be obtained by videotaping the copiers and having a limited number of judges code the tapes according to an agreed scheme. There are, however, two biases in the text-copying results: as mentioned earlier, the constituency of the beginnings of sentences probably tends to be blurred on account of relatively few pauses being attracted there, and very short constituents are probably not given due prominence. One approach to these problems would be to simply ignore those parts of the constituent bracketings where the biases are *a priori* likely to have operated. Another solution would be to attempt to estimate the magnitude of the biases and to weight the pause distributions accordingly.

Footnotes

1 Various parts of this work have formed the basis of talks given in a number of universities in Britain and India, as well as a short paper presented to the 1982 Summer Meeting of the Linguistic Society of America. I am grateful for audience comment, which has been most helpful on all occasions.

2 I am inclined to add to the desiderata of naturalness of task, simultaneity and sensitivity: that an ideal psycholinguistic method for the elucidation of constituent structure should also be economical in the use it makes of subjects' and experimenters' time. Some may feel that economy is an extraneous consideration in 'pure science', but an uneconomical technique is unlikely to be used outside of feasibility demonstrations, whereas one which easily gives information on many sentences stands a chance of becoming a real tool of discovery. Of the methods mentioned in the text, Levelt's, Johnson's and the click approach seem to me to be uneconomical ones. For instance, for a seven-word sentence (which might be uttered in under two seconds) Levelt would require considered judgements on 21 pairs of words from each subject. The pause studies, to be discussed next, make better use of subjects' time, since a small corpus of recorded speech will contain large quantities of data (though analysis is bound to be time-consuming).

3 The idea of the text-copying method was suggested to me by a technique which Chase & Simon (1973) used to investigate chess playing.

4 Grosjean & Lane (1977) independently devised what amounts to the same algorithm.

5 Height in the tree could be scaled to reflect the number of copiers pausing at each boundary, but I have chosen not to do so in this article. Such a scaling has no obvious theoretical linguistic interpretation, though it might be a useful graphic aid for suggesting whether or not a given bracketing is reliably distinct from alternative possibilities.

6 The data from this set of five-word sentences provide a convenient illustration of the algorithm, but it will be noted in the account given later of Experiment 2 that the small number of sentence-internal pauses involved makes this particular result an unreliable one.

7 The line of argument presented in the body of the paper gives the benefit of doubt to those who would

posit word-length as the determining factor responsible for the Figure 3 results. One could, conversely, argue that pauses should be expected immediately after short words: subjects should stop reading and empty their short-term memory contents via their pens before tackling a long word; then, having subsequently absorbed the long word, they will often find that they still have room for one or more short words. This strategy would yield tree diagrams that mostly violated theoretical linguistic accounts of English sentence structure.

8 See Vestergaard (1977) for a monograph-length study on this topic and closely related matters.

9 For some discussion of alternative anlyses of relative clauses, see Stockwell, Schachter & Partee (1973: ch 7) and, especially, Partee (1976: 53f).

10 An unanticipated third interpretation, 'draft (a picture of) the body of the tongue', was offered by some of the judges who (see below) were subsequently asked to paraphrase the ambiguous sentences. This reading is also presumably related to a phrasal verb analysis of draw up.

11 I am grateful to John Local and Kit-Ken Loke for assistance in conducting Experiment 2.

12 X^2 'goodness of fit' comparisons are rendered infeasible here by the numbers of pauses in some of the relevant loci being too small.

13 Here, and where they are reported elsewhere below, average rank coefficients, $\bar{r}$', were derived from coefficients of concordance corrected for continuity and, when necessary, for ties.

14 It should be pointed out that the small numbers of pauses which fell in some of the inter-word gaps made it necessary to combine some adjacent clusters of pauses for the purpose of calculating the X^2 values presented in Table 8. This reduced the degrees of freedom to 6, from a potential 11. The sequences within which pauses were combined and treated as if they had occurred in a single location were the same for all three types of sentence: the initial AdvP, the string Aux-to-V and the sentence-final PP-modified NP (an example of the latter being the shape of each organ). Even so, one of the 'expected' frequencies was an uncomfortably low 3.8. This slightly exaggerates the resulting X^2 value, but as it occurs in the comparison between PrV and VPP - the non-significant X^2 - it simply adds force to the argument that the PrV and VPP sentences have the same structure.

References

Bever, T G, J R Lackner and R Kirk (1969) 'The underlying structures of sentences are the primary units of immediate speech processing.' Perception and Psychophysics, 5, 225-234.

Butterworth, B (1980) 'Evidence from pauses in speech'. In B Butterworth (ed) Language production. Vol 1: Speech and talk. London: Academic Press, 155-176.

Chase, W G and H A Simon (1973) 'The mind's eye in chess'. In W G Chase (ed) Visual information processing. New York: Academic Press, 215-281.

Frazier, L and J D Fodor (1978) 'The sausage Machine: a new two-stage parsing model'. Cognition, 6, 291-325.

Fodor, J A, T G Bever and M F Garrett (1974) The psychology of language. New York: McGraw-Hill.

Grosjean, F and H Lane (1977) 'Pauses and syntax in American Sign Language'. Cognition, 5, 101-117.

Grosjean, F, L Grosjean and H Lane (1979) 'The patterns of silence: performance structures in sentence production'. Cognitive Psychology, 11, 58-81.

Halliday, M A K (1976) 'The English verbal group'. In G Kress (ed) Halliday: system and function in language. London: Oxford University Press, 136-158.

Hawkins P R (1971) 'The syntactic location of hesitation pauses'. Language and Speech, 14, 277-288.

Huddleston, R (1976) An introduction to English transformational syntax. London: Longman.

Johnson, N F (1965) 'Linguistic models and functional units of language behavior'. In S Rosenberg (ed) Directions in psycholinguistics. New York: Macmillan, 29-65.

Kimball, J (1973) 'Seven principles of surface structure parsing in natural language'. Cognition, 2, 15-47.

Larsen, S F (1971) 'The psychological reality of linguistic segments reconsidered'. Scandinavian Journal of Psychology, 12, 113-118.

Levelt, W J M (1969) 'The scaling of syntactic relatedness: a new method in psycholinguistic research'. Psychonomic Science, 17, 351-352.

Levelt, W J M (1970) 'A scaling approach to the study of syntactic relations'. In G B Flores d'Arcais and W J M Levelt (eds) Advances in psycholinguistics. Amsterdam: North Holland, 109-121.

Levelt, W J M (1978) 'A survey of studies in sentence perception: 1970-1976'. In W J M Levelt and G B Flores d'Arcais (eds) Studies in the perception of language. Chichester: Wiley, 1-74.

Martin, E (1970) 'Toward an analysis of subjective phrase structure'. Psychological Bulletin, 74, 153-166.

Olson, G M and H H Clark (1976) 'Research methods in psycholinguistics'. In E C Carterette and M P Friedman (eds) Handbook of perception. Vol VII: Language and Speech. New York: Academic Press, 25-74.

Partee, B H (1976) 'Some transformational extensions of Montague grammar'. In B H Partee (ed) Montague grammar. New York: Academic Press, 51-76.

Quirk, R, S Greenbaum, G Leech and J Svartvik (1972) A grammar of contemporary English. London: Longman.

Stockwell, R P, P Schachter and B H Partee (1973) The major syntactic structures of English. New York: Holt, Rinehart and Winston.

Vestergaard, T (1977) Prepositional phrases and prepositional verbs. The Hague: Mouton.

Zwicky, A M (1978) 'Arguing for constituents'. Chicago Linguistic Society: Proceedings of the 14th Regional Meeting. 503-512.